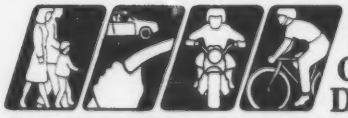
Accident Potential: An Ontario Driver Records Study Summary Report

A. Smiley
Human Factors North

E. Hauer
B. Persaud
Department of Civil Engineering
University of Toronto

L. Clifford
D. Duncan
Ministry of Transportation

April, 1991
ISBN 0-7729-5847-5
SCDO 91-115



Safety Co-ordination & Development Office
Research & Evaluation Section

Ministry of Transportation
Ontario

92-00131/

ABSTRACT

The current demerit point system in use in Ontario allocates points to offences on the basis of the perceived seriousness of the offence. The goal of the work reported here was to allocate points to offences with a different purpose in mind. The purpose here was to use a driver's record of convictions and accidents in order to predict, as well as possible, which drivers, based on their past record, are most likely to have an accident in the near future.

A sample of 827,955 records of Ontario drivers containing information about age, gender, convictions, accidents, demerit points and suspensions, for 1981-1984 was examined. On this basis 16 new models to estimate a driver's accident potential were formulated. Some models used age and gender information, others did not. Some models assigned different weights to each of the conviction categories; in other models all convictions were weighted equally (equivalent to 1 point per conviction). In some models accident data were used; in other models, not. In contrast to the current system, information on all types of convictions, not just moving convictions, was used. It turns out that these new models are better at predicting a driver's accident potential than the current demerit point system.

The performance of the new models is described by means of two measures. The first measure is the number of accidents in period 2 (1983-1984) based on the driver's record of convictions and accidents in period 1 (1981-1982). For example, if the 10,000 drivers with the **most demerit points** in period 1 are selected, 1452 have accidents in period 2; if the 10,000 drivers with the **most accidents** in period 1 are selected, 1828 have accidents in period 2. Alternatively, if the 10,000 drivers are selected on the basis of the **best new model**, which uses age, gender, accidents and separate weights for 14 offence categories, 2116 have accidents in period 2.

The second measure of performance is in terms of "hits" and "false alarms". "Hits" are drivers who have an accident potential in excess of 4 times the population average. For example, if the 10,000 drivers with the highest estimated accident potential are selected by the best model, 3757 of these are are expected to be "hits". "False alarms" are drivers who have an accident potential which is below the population average. Of the same 10,000, one should expect 792 to be "false alarms". The best model identifies approximately twice as many high accident potential drivers as the current demerit point system. Even the simplest model, which uses total convictions as the only variable, predicts 50% more high accident potential drivers than the current system.

ABRÉGÉ

L'actuel système de points d'inaptitude en usage en Ontario donne des points aux infractions en fonction de la gravité qu'on leur attribue. Dans le travail qui nous occupe ici, il s'agissait de donner des points aux infractions dans un but différent, plus précisément, de se baser sur le dossier des condamnations et accidents des conducteurs pour prédire le mieux possible lesquels ont le plus de chances d'avoir un accident dans un avenir proche.

On a donc examiné un échantillon de 827 955 dossiers de conducteurs ontariens, contenant des renseignements sur l'âge, le sexe, les condamnations, les accidents, les points d'inaptitude et les suspensions de permis de 1981 à 1984. À partir de ces données, on a établi 16 nouveaux modèles censés évaluer les chances d'accident d'un conducteur. Certains modèles intégraient les données sur l'âge et le sexe, d'autres pas. Certains modèles pondéraient les diverses catégories de condamnations et d'autres les traitaient toutes sur le même pied (un point par condamnation, quelle que soit la nature du délit). Certains modèles intégraient les données sur les accidents et d'autres pas. Contrairement au système actuel, ces modèles tenaient compte de tous les types de condamnations, et non uniquement de celles portant sur des infractions commises par les conducteurs de véhicules en marche. Il s'est avéré que ces nouveaux modèles arrivent à mieux prédire les chances d'accident d'un conducteur que l'actuel système de points d'inaptitude.

Les résultats des nouveaux modèles apparaissent dans deux mesures. La première est le nombre d'accidents pendant la période 2 (1983-1984) en fonction du dossier des condamnations et accidents du conducteur pendant la période 1 (1981-1982). Par exemple, si l'on retient les 10 000 conducteurs qui ont eu le plus de points d'inaptitude pendant la période 1, on trouve que 1 452 d'entre eux ont des accidents pendant la période 2; si l'on retient les 10 000 conducteurs qui ont eu le plus d'accidents pendant la période 1, on trouve que 1 828 d'entre eux ont des accidents pendant la période 2. Selon l'autre méthode, si la sélection des 10 000 conducteurs s'effectue selon le meilleur nouveau modèle, basé sur l'âge, le sexe, les accidents et 14 catégories d'infractions de poids différents, on trouve que 2 116 d'entre eux ont des accidents pendant la période 2.

La deuxième mesure classe les conducteurs en «récidivistes» et «fausses alertes». Les récidivistes sont les conducteurs qui ont des chances d'accident supérieures à quatre fois celles de la population moyenne. Par exemple, si les 10 000 conducteurs qui ont le plus de chances d'accident sont sélectionnés par le meilleur modèle, 3 757 d'entre eux sont classés comme «récidivistes». Les «fausses alertes» sont les conducteurs qui ont des chances d'accident inférieures à celles de la population moyenne. Sur les 10 000 conducteurs cités ci-dessus, on devrait en trouver 792. Le meilleur modèle permet d'identifier environ deux fois plus de conducteurs ayant des chances élevées d'accident que l'actuel système de points d'inaptitude. Même le modèle le plus simple, dont la seule variable est le nombre total de condamnations, identifie 50 % de plus de conducteurs ayant des chances élevées d'accident que l'actuel système.

Table of Contents

1.0	INTRODUCTION	Page 1
2.0	DRIVER RECORD SAMPLE	1
3.0	PREPARATION FOR ANALYSIS	2
3.1	Checks for Representativeness of the Sample	2
3.2	Consistency Checks	2
3.3	Exploratory Analysis	3
3.4	Selection of Conviction Categories	3
3.5	Discussion of Conviction Categories	11
3.6	Age and Gender Categories	12
3.7	Exclusion of Drivers With Suspended Licenses	12
4.0	ANALYSIS AND RESULTS	14
4.1	Schemes and Variants Examined	14
4.2	How a Model Works	16
4.3	The Accident Potential of Ontario Drivers	16
4.4	The Performance of the Current Demerit Point System and of the New Models	18
5.0	SUMMARY AND CONCLUSIONS	30

Figures and Tables

Figure 1	4 Year Convictions Vs. 4 Year Accidents	Page 4
Figure 2	2 Year Convictions Vs. 2 Year Accidents	5
Figure 3	Expected Numbers of Accidents in a Subsequent 3-Year Period for Drivers with 1 Conviction in Any Year	10
Figure 4	Relationship Between Age and Accidents	13
Figure 5	Distribution of Accident Potential in the Ontario Driver Population	19
Figure 6	Expected Accidents/year	24
Figure 7	Current DP Vs. Schemes A4, B4, C4, D4	28
Figure 8	Runs A4, B4, C4, D4	29
Table 1	Accidents for Drivers with 1 Conviction in 1 Year	7
Table 2	Disaggregation of Speeding Convictions	8
Table 3	3-Yr. Mean Accidents for Drivers with 1 Conviction in 1 Year	9
Table 4	Variables Used for Regression Runs	15
Table 5	Regression Weights for Model A2	17
Table 6	Accident Per Year Recorded by Drivers Selected by Various Models	21
Table 7	Figures of Merit-A4, B4	23
Table 8	Figures of Merit for 10000 Drivers With Highest Accident Potential	27

1.0 INTRODUCTION

The current demerit point system in use in Ontario allocates points to offences on the basis of the perceived seriousness of the offence. An offence is considered serious if it is associated with a relatively large chance of precipitating an accident. This is why a non-moving violation, such as not having a trailer permit, receives no points, while running a red light receives a number of points. The goal of the work described in this report was to allocate points to offences with a different purpose in mind. The purpose here was to use a driver's record of convictions and accidents in order to predict, as well as possible, which drivers, based on their past record, are most likely to have an accident in the near future. This is determined by calculating, for each driver, how many accidents per year a driver is likely to have, on the average.

Because a person's "accident potential" can only be indirectly estimated (not directly measured) and because, mercifully, for any one driver it is rare to be involved in an accident, the accuracy with which we can estimate the accident potential of a driver is bound to be severely circumscribed. Thus, our aim is not only to produce for each driver an estimate of his or her accident potential, but also to say how accurate our estimate is.

This is the kind of knowledge which might then be used in the determination of post-licensing control action. Thus while a non-moving violation may not be a threat to traffic safety, such a conviction on a driver's record may be an important clue about that person's likelihood of future accidents.

2.0 DRIVER RECORD SAMPLE

The analysis examined driver records over a recent four year period. Convictions and accidents during the first two years of the record were used as the basis for predicting accidents in the second two year period.

To be included in the sample, drivers had to have been licensed to drive in Ontario throughout the four year period. To this end, a number of criteria were used in the selection of driver records. These criteria were as follows. First, drivers not licensed at the beginning of the study period were excluded. Second, drivers who moved out of Ontario during the four year period were excluded. Third, drivers who had ceased driving during the study period were excluded. These exclusions were achieved by only selecting drivers who renewed their drivers' license during 1985, with the exception of those who were 16 or 17 in 1981. This exception was made because such a large percentage of drivers become licensed at 16 and 17, that the majority of drivers of this age would be renewing their licenses in 1984 (3 years after original licensing) and would have been excluded from the sample.

Of the 5.5 million Ontario drivers, 827,995 qualified for inclusion in the sample. Driver record data extracted from the MTO driver records included the following information: age, gender, and, for each conviction: type, date and demerit points assigned, for each accident: degree of severity, date, and for each suspension: type and time period. The processing of this amount of data can be costly. For some purposes, smaller data sets are

sufficient. With this in mind, three sets of driver records were provided: a full set of records, and smaller samples of every 40th record and every 400th record.

3.0 PREPARATION FOR ANALYSIS

Making sense of large data sets requires careful preparation. The several preparatory activities are described below.

3.1 Checks for Representativeness of the Sample

Several tests were conducted to check whether the characteristics of a smaller sample (1 in 40 records) were similar to known (published) characteristics of the Ontario driving population. In terms of percent of male and female drivers, percent of accidents involving males, and accident involvement per driver/year the sample data were found to be reasonably representative of the Ontario driving population.

3.2 Consistency Checks

Driver record data were then checked for consistency. A number of inconsistencies were noted. The first inconsistency apparent in the data was that while some offences are logically associated with the occurrence of an accident, the count of such offences exceeded the count of related accidents. Thus, for example, one would expect that if a driver were convicted of failing to remain at the scene of an accident, he or she would also show an accident associated with the offence on that same day. However, this is clearly not so; in the small sample there were 45 such convictions, but only 28 accidents for drivers convicted of this offence. There is a similar discrepancy for convictions for "following too close" which apparently arise mainly from accidents. These discrepancies suggest that many accidents related to convictions have not been entered in the driver record file. It is possible that this omission might apply to accidents in general. One reason for this may be due to the fact that records contain only reports of "reportable" accidents whereas charges apply to any.

A second inconsistency was the considerable year to year fluctuation in the number of convictions for certain offences. For example, the number of convictions for driving without a permit increased from 0 in 1981 to 68 in 1984. There were seven other conviction types with similar inconsistencies from year to year. The reason for these inconsistencies is not clear, but they could be related to changes in enforcement practices or modifications related to regulations.

The third inconsistency involved the original licensing date. This variable was found to be unreliable. Some of the dates were missing and none of the original dates were earlier than 1966 although a substantial number of drivers were obviously licensed before that year. This meant that the effect of experience on a driver's accident record could not be examined. This is a result of the development of the computer system in 1966 and of changes of residence and lapses of renewal that affect the recording of the original date.

3.3 Exploratory Analysis

The sample of every 40th record was used to perform some exploratory analyses. Figure 1 shows the relationship between accidents and convictions in four years in a sample of nearly 16,000 drivers. It shows that drivers with more convictions tend to have more accidents. Thus, for example, the 189 drivers with 6 convictions in four years had, on the average, eight times as many accidents (line 1) as drivers with no convictions in these four years. (Convictions which were incurred on the date of an accident were eliminated from consideration). The same pattern prevails for at-fault accidents (line 2), fatal and injury accidents (line 3) and at-fault, fatal, and injury accidents(line 4). Figure 2 shows similar results. However, this time the relationship between convictions in the first two-year period and the average number of accidents in the subsequent two-year period is plotted.

These two figures leave no doubt that convictions and accidents are correlated. However, correlation should not be taken to indicate causality. It is in fact very likely that the number of accidents in which a driver is involved and the number of convictions which a driver incurs are both related to the distance driven. Thus, the observed statistical correlation reflects a hidden common link, exposure, about which we have no information.

3.4 Selection of Conviction Categories.

A preliminary analysis of a sample of about 8000 drivers indicated that, over the four-year period 1981-1984, these drivers were convicted for approximately 200 different traffic offences. Speeding accounted for some 60% of non-accident related convictions; seat-belt offences accounted for about 10%, and failure to stop at an intersection for about 5%. Most of the other offence types had very few convictions. It was obviously impractical to assign different weights for each of 200 conviction types. Nor was it feasible to obtain statistically reliable estimates of how much these "leaner" convictions added to a driver's expected number of accidents. It was therefore necessary to place those offences with few convictions into larger groups. (It should be noted that the same approach is already present in the current point-system. There are essentially 7 categories of offences, that is, those assigned 0,2,3,4,5,6 and 7 demerit points. For example, all the non-moving violations as well as some of the moving ones fall into the 0 point category.)

Determining Conviction Categories: Step 1. The clustering of all possible offence types into a manageable number of groups was accomplished in several steps. The first step was to combine the offence types which are quite similar in nature into a smaller number of categories. A total of 45 such categories were established in our discussions with MTO in April 1987. These consisted of 38 categories of moving violations, 3 for vehicle-related offences, and 4 for non-moving, administrative violations.

4 YRS CONVICTIONS VS 4 YRS ACCIDENTS

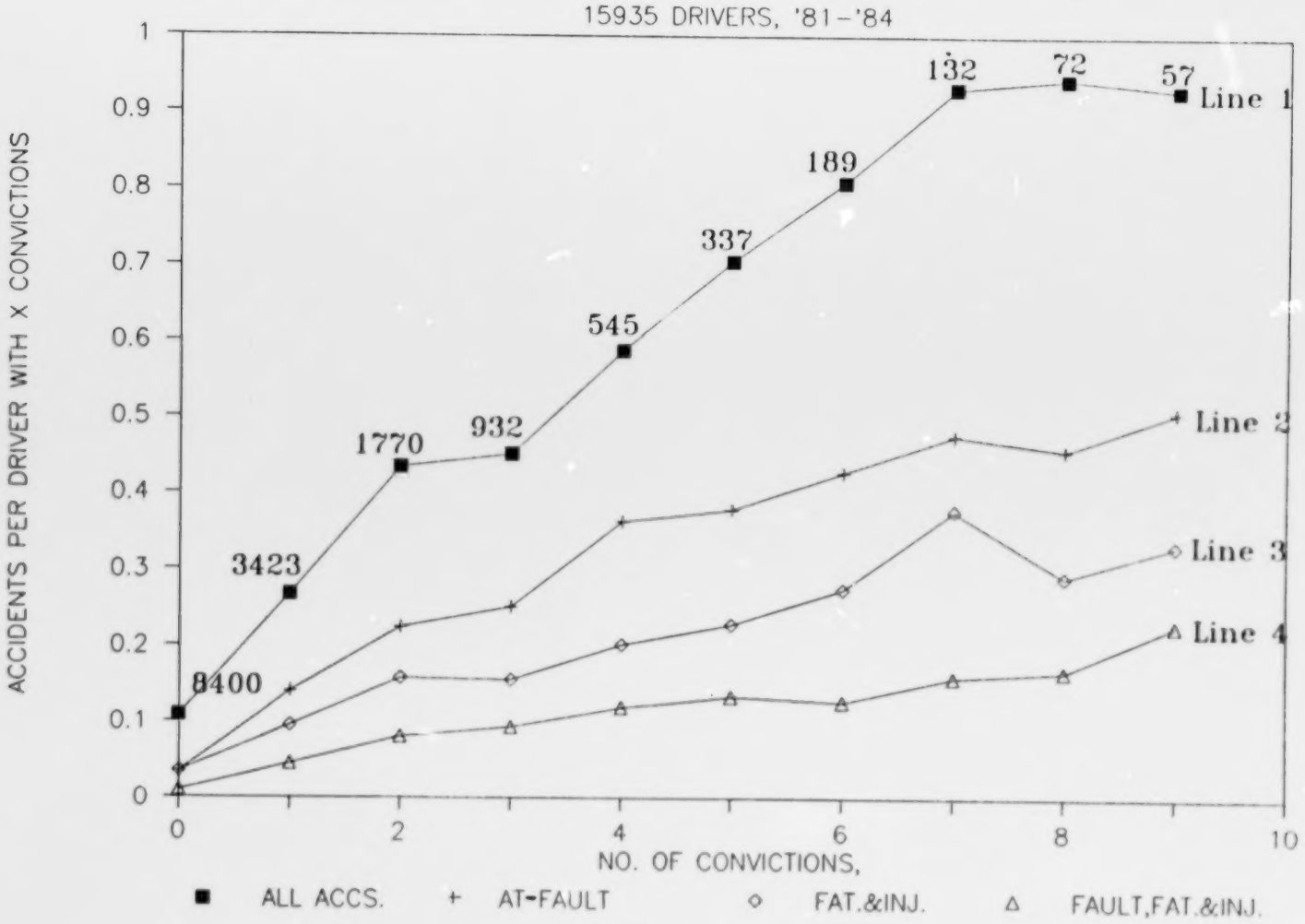


Figure 1

2 YRS CONVICTIONS VS 2 YRS ACCIDENTS

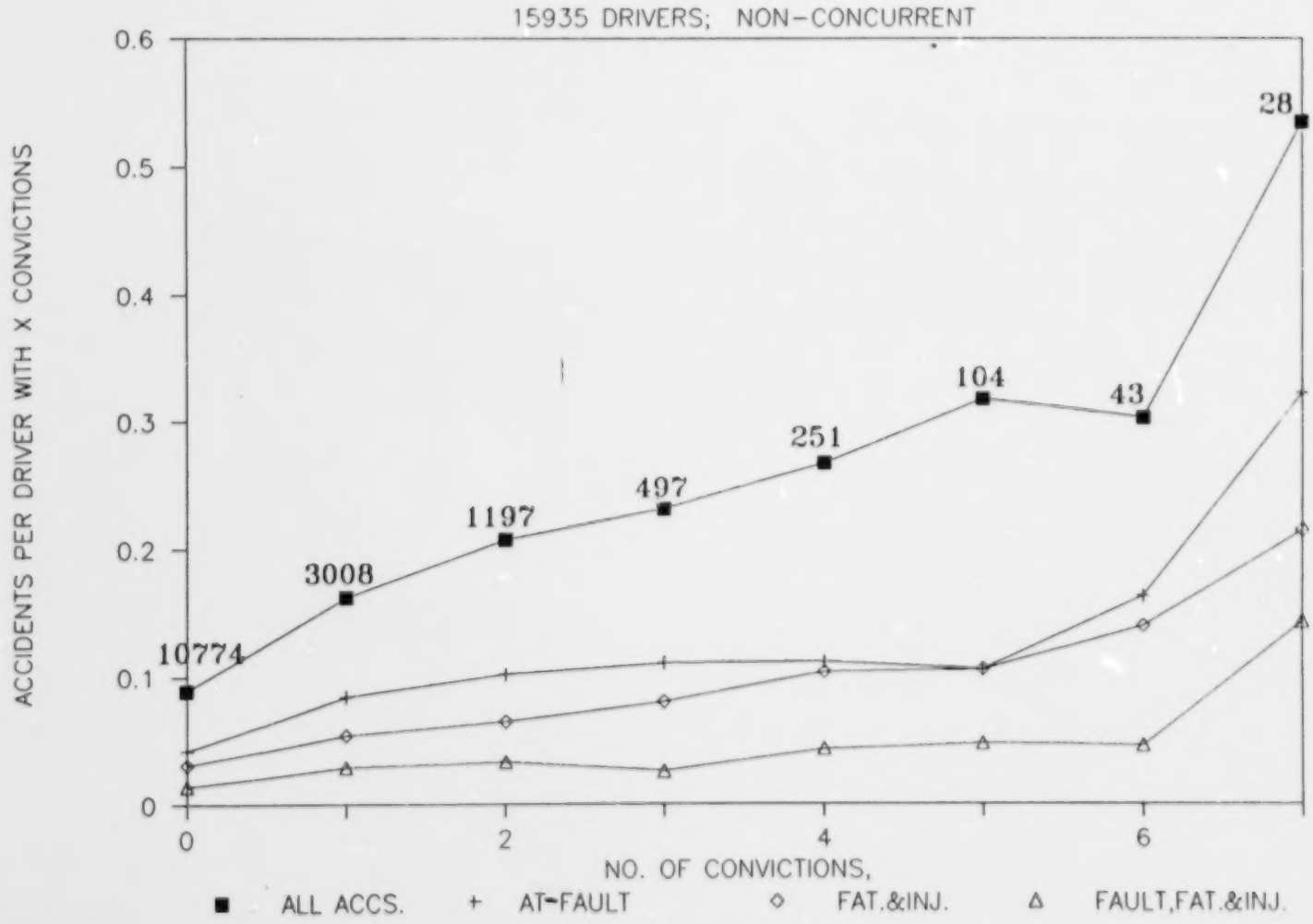


Figure 2

The first round of clustering was based on judgment of similarity. In the second round, groups of offences were identified which were similar in their contribution to the average number of accidents of a driver. To obtain results which are unambiguous and free of confounding, the only records used were those for drivers who, in any one of the four years, had a single conviction. The results are given in Table 1 and are based on the entire sample of drivers with a single conviction in four years. If, at this stage, records had been used containing more than 1 conviction, it would have been impossible to know how much each conviction was related to the probability of a subsequent accident. Thus, for example, there were 12337 drivers who in one year were convicted for not wearing a seat belt (offence m1) and had no other conviction in that year. During the remaining three years these drivers recorded 4858 accidents for an average of 0.394 accidents per driver. To eliminate any bias due to differences in the age-gender distribution (i.e. number of male-female, young-old drivers in each conviction group) which might be associated with specific offences, all averages were recalculated for a "standard population". The "standard population" used had an age-gender composition which was characteristic of those drivers who had exactly one conviction of any kind in one of the four years. This is why in the "Weighted Mean" column the average number of accidents in the remaining three years associated with this seat belt offence is listed as 0.376 rather than 0.394. Similarly, the 173,592 drivers who had only a single speeding conviction in some year have an adjusted average of 0.319 accidents in three years. The last two columns give 95% confidence limits for the weighted mean. The confidence limit measures the accuracy with which the mean is estimated. The narrower the confidence limit, the more accurate the estimate.

Speeding. The current point allocation system establishes four levels of severity according to how much the speed limit is exceeded. However, exceeding the speed limit by 20 km/h, whether in a 30 km/h or a 100 km/h speed zone, carries the same amount of demerit points. Since there were many speeding convictions in the data base, the question of how the average number of accidents depends on the speed limit and on the extent to which the speed limit is exceeded was examined. The results are given in Table 2. It turns out that there is no systematic or practically significant difference between the average number of accidents either by speed limit zone or by the amount by which the speed limit is exceeded, or the combination of both criteria.

Determining Conviction Categories: Step 2. On the basis of such comparisons, it was possible to combine some of the 45 conviction categories which were found to be associated with a similar average number of accidents. The resulting 14 conviction groupings, the associated three-year average number of accidents and 95% confidence limits of the estimate are shown in Table 3 and plotted in Figure 3. Also shown are estimates of accident potential for conviction-free drivers -- those who had no convictions of any type during one calendar year.

TABLE 1: Accidents for Drivers with 1 Conviction in 1 Year

Category	Brief Description	No. of Drivers	3 yr. Accs.	Weighted Mean	95% Limits Upper	Lower
n1	Minor neglect,licenses,permits	6495	2918	0.434	0.445	0.422
n2	Neglect,insurance,permits,etc.	1589	719	0.414	0.438	0.392
n3	License suspended,HTA	874	454	0.424	0.456	0.394
n4	Learners	34	18	0.343	0.502	0.212
v1	Minor veh.;lamps,noise	2954	1498	0.468	0.485	0.451
v2	Brakes, tires,unsafe vehicle	946	451	0.400	0.430	0.371
v3	Comm.veh;size & weights	503	369	0.542	0.583	0.500
m1	Seat belt	12337	4858	0.376	0.384	0.368
m2	Speeding	173592	55211	0.319	0.321	0.317
m3	Careless driving	902	342	0.327	0.357	0.299
m4	Slow driving	45	11	0.119	0.237	0.055
m8	STOP sign, ROW violations	14024	3935	0.288	0.295	0.281
m9	PXO violations	1237	355	0.296	0.320	0.272
m10	Turns violations;right,left,U	18231	4942	0.283	0.289	0.277
m11	Unsafe move;open door	1649	542	0.334	0.355	0.312
m13	Disobey red light	13731	4270	0.313	0.321	0.306
m14	Amber light	3453	982	0.285	0.299	0.271
m15	Advance green	274	73	0.265	0.317	0.218
m16	Fail to share road	170	62	0.303	0.372	0.242
m17	Passing violations	1305	459	0.327	0.351	0.303
m18	Wrong-way one way street	1582	458	0.284	0.306	0.264
m19	Improper driving divided h'way	2599	900	0.361	0.379	0.344
m20	F.T.C.	934	337	0.344	0.374	0.316
m21	Emerg. veh., school x'ing	48	15	0.159	0.280	0.084
m22	R/R crossing violations	95	35	0.314	0.408	0.233
m24	Headlight beam not lowered	225	71	0.260	0.318	0.209
m25	Improper parking	145	77	0.407	0.484	0.334
m26	Fail stop for school bus	604	133	0.281	0.317	0.249
m28	Disobey traffic signs	1650	529	0.322	0.344	0.301
m29	Fail report accident	224	73	0.266	0.324	0.215
m30	Fail remain at scene	236	75	0.315	0.373	0.261
m32	Dangerous driving C.C.C.	5	2	0.010	0.421	0.000
m33	Fail remain at accident C.C.C.	66	39	0.377	0.491	0.274
m34	Dangerous driving C.C.C.	89	46	0.281	0.376	0.202
m35	Impaired driving C.C.C.	2381	1040	0.443	0.462	0.424
m36	Fail/refuse breath test C.C.C.	94	37	0.218	0.306	0.149
m37	Fail or ref. breath test C.C.C.	120	50	0.314	0.397	0.241
m38	Driving with >80 mgs. alcohol	3676	1502	0.386	0.401	0.371
m41	Crowding driver seat	120	36	0.280	0.362	0.211
m44	Radar device in vehicle	68	27	0.251	0.359	0.166
m45	No safe helmet, motorcycle	96	47	0.266	0.357	0.191
m46	Fail to signal to stop	18	6	0.080	0.277	0.019
m47	FTC, commercial vehicle	67	35	0.362	0.475	0.262
m48	Fail to stop for police officer	12	6	0.086	0.340	0.016

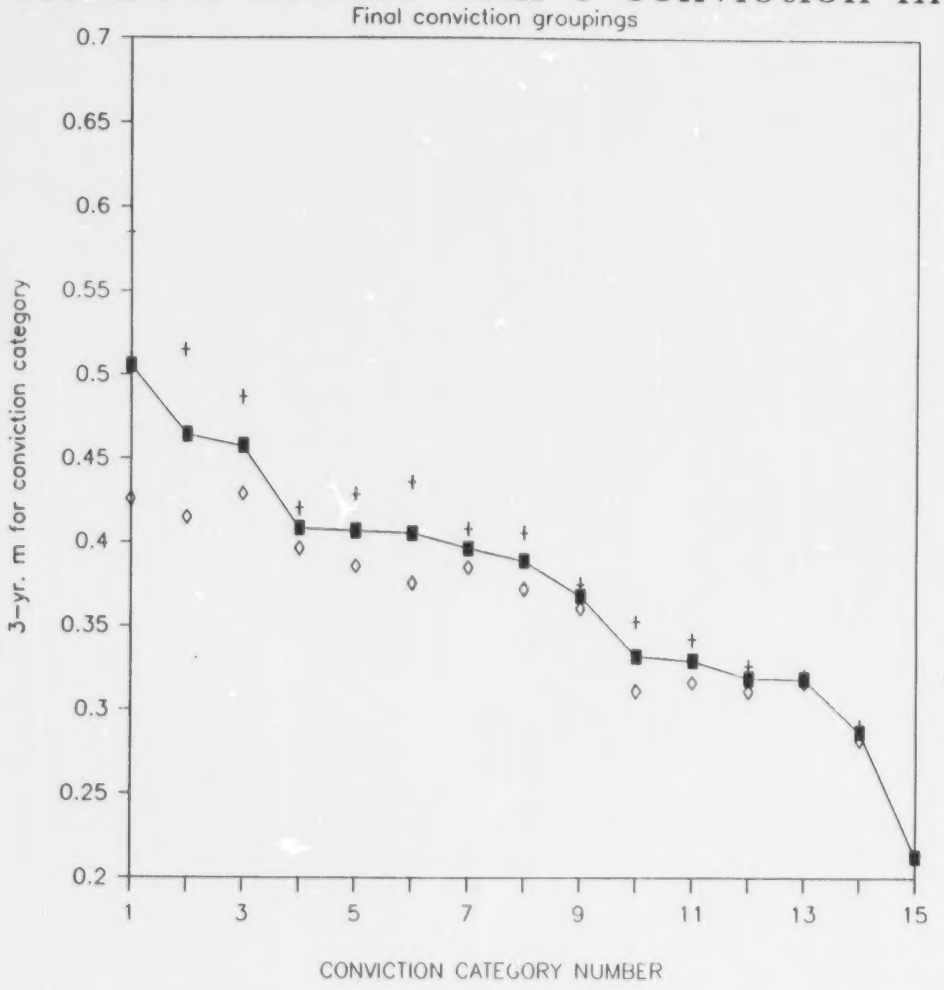
TABLE 2: Disaggregation of Speeding Convictions

Speed Limit	Above Limit	No. of Drivers	3 yr. Accs.	Weighted Mean	95% Limits Upper	Lower
All	0-15	123224	33755	0.293	0.296	0.291
All	16-29	79589	22479	0.286	0.289	0.283
All	30-49	23891	7331	0.297	0.302	0.291
All	>49	894	371	0.315	0.344	0.286
20-40	0-15	24264	6307	0.288	0.293	0.283
50-70	0-15	80243	22037	0.294	0.297	0.291
80-100	0-15	18712	5411	0.292	0.298	0.286
20-40	16-29	5276	1413	0.281	0.293	0.270
50-70	16-29	39995	11427	0.293	0.297	0.288
80-100	16-29	34318	9639	0.281	0.285	0.276
20-40	30-49	683	185	0.269	0.302	0.239
50-70	30-49	9304	2914	0.305	0.314	0.296
80-100	30-49	13904	4232	0.292	0.300	0.285
20-40	>49	14	8	0.186	0.432	0.064
50-70	>49	405	183	0.341	0.386	0.299
80-100	>49	475	180	0.272	0.312	0.236

TABLE 3. : 3-YEAR MEAN ACCIDENTS FOR DRIVERS WITH 1 CONVICTION IN 1 YEAR. (Based on Final Conviction Categories)

Category (Final)	Brief Description	Drivers	Accidents	3-Year Mean	90% Limits Upper	Lower
mm1	Fail to yield, improp.turns,etc	37775	10425	0.287	0.291	0.282
mm2	Disobey red lights etc.	13183	4169	0.319	0.326	0.311
mm3	Fail report acc; C'less driving	1774	657	0.332	0.353	0.311
mm4	Unsafe move, overtaking	4606	1514	0.329	0.342	0.316
mm5	Fail to remain; breath test	6263	2563	0.396	0.408	0.385
mm6	Seat belt; FTC; parking	15454	5885	0.367	0.375	0.360
mm7	Speeding	165852	52623	0.318	0.321	0.316
nn1	Minor neglect;plates,license	2865	1166	0.388	0.405	0.372
nn2	Neglect;d/lic. suspended	5720	2485	0.408	0.420	0.396
vv1	Minor vehicle:lamps, etc.	1802	786	0.406	0.428	0.385
vv2	Vehicle:tires,brakes, unsafe	903	441	0.405	0.436	0.375
vv3	Insecure Load	134	94	0.505	0.584	0.425
vv4	Truck weight/dimension offences	348	214	0.464	0.513	0.415
ee1	Environmental offences	1022	527	0.457	0.486	0.428
	Drivers with no convictions	2047414	337745	0.216	0.213	0.212

Expected Numbers of Accidents in a Subsequent 3-Year Period for Drivers with 1 Conviction in Any Year



15 No conviction group
14 mm1: Fail to yield, imp. turns, PXO,
13 mm7: Speeding offences
12 mm2: Disobey red lights; etc
11 mm4: Unsafe move;imp. o'taking; etc.
10 mm3: Fail report accid.; Careless driving;
9 mm6: Seat belt; F.T.C.; parking;
8 nn1: Minor neglect: no drivers license;
7 mm5: Fail to remain; breath test;
6 vv2: Vehicle neglect: Unsafe veh.,
5 vv1: Minor vehicle neglect: lamps
4 nn2: Neglect: D/Lic. suspended etc.
3 ee1: Environmental offences:
2 vv4: Truck weight/dimension offences.
1 vv3: Insecure Load.

+ Upper Confidence Level

◊ Lower Confidence Level

Figure 3

In summary, the final 14 conviction categories to be used in analysis were established on the basis of the following considerations:

1. Conviction types within each category were similar in nature;

2. The accident potential associated with each conviction within a category was similar in size;

3. The numbers of drivers having offences in each category was sufficient to provide a reliable estimate of accident potential for that category.

3.5 Discussion of Conviction Categories

Inspection of Figure 3 leads to the question of why it is that convictions for some of the innocuous offences (e.g. for a burnt-out light bulb) are found to be associated with more accidents than those offences traditionally deemed very dangerous, such as speeding or running a red light. Several reasons combine to explain this apparent paradox. First, convictions associated with most accidents are those characteristic of truck drivers. These drivers cover 10-20 times the distance of a passenger car driver. Therefore, it is to be expected that they will have, on the average, more accidents.

The second reason is easiest to explain through an example. Assume that 1000 run-the-red offences lead to 5 accidents and 1000 fail-to-signal-turn offences lead to 1 accident. Thus, we are assuming that to run the red is a more dangerous offence than failing to indicate a turn. However, the enforcement of both offences is unequal (perhaps because one is assumed to be more dangerous than the other). Assume that of the 1000 run-the-red offences, 10 lead to a conviction, whereas of the 1000 fail-to-signal offences, 1 leads to a conviction. Thus, in a figure similar to Figure 3, that would be 5/10=0.5 accidents per conviction for running the red and 1/1=1.0 accidents per conviction for failure to indicate a turn. Even though, according to our starting assumption, running the red is five times more dangerous than failing to indicate a turn, because enforcement of the two offences is unequal, the final results show the contrary. The problem is caused by the fact that the driver record contains information about convictions, not the number of illegal actions committed by a driver for which he was not convicted. This is the data with which we have to work.

A third reason may be related to the connection between different types of behaviour and convictions. The incidents on a person's driving record, convictions and accidents, might be indications of their overall driving behaviour. As a result, the types of convictions committed by certain types of people may also provide insight into their potential for accidents. If a person engages in certain behaviours which lead to certain convictions, they may also engage in certain other behaviours which predispose them to accidents. To illustrate with an example, drivers with environmental types of convictions (e.g. noisy muffler) were found to have a higher weighted mean of accidents than drivers with most other types of conviction. Most drivers quickly have their noisy muffler fixed and are unlikely to receive this type of conviction. The attitude which results in drivers coming to the attention of the police and being charged with this offense may be related to a similar careless attitude towards behaviour which results in accidents.

All of the three reasons discussed above affect the correlation between a particular offense category and accidents. Thus, correlation does not necessarily indicate causality. Two conclusions follow. First, one should not interpret the results in Figure 3 as providing information about the danger inherent in this or that offence. Second, one should not be surprised when, later, in section 4, the same innocuous offences will prove to be strongly related to the driver's estimated accident potential.

3.6 Age and Gender Categories

It is well known that the average number of accidents for a driver depends on gender and age. To account for this fact, age and gender will be used in the analysis as "variables". It is relatively simple to account for gender since it comes in two natural categories. However, the relationship between age and number of accidents is continuous in nature and distinctly non-linear as is shown in Figure 4. To include age in the analysis, it was necessary to establish a number of age categories. After careful analysis, the boundaries between age groups were chosen so that the average numbers of accidents within each group remained nearly constant while still keeping sufficient numbers of drivers within each age category to maintain statistical reliability. The following eight age categories were chosen:

< 21
21-25
26-30
31-35
36-40
41-50
51-60
> 60

3.7 Exclusion of Drivers with Suspended Licenses

Some drivers in the sample had their license suspended during the study period. Most of these are drivers who had many convictions which carried points. The extent to which a suspended driver curtails his or her driving is unknown. This leads to serious difficulty in the statistical analysis. Consider a driver who in the first two-year period had many convictions and was suspended. That driver can be expected to drive less in the second two-year period and therefore to have proportionately fewer accidents. In the statistical analysis this would tend to create a negative correlation; that is, it would lead to the incorrect result that the larger the number of convictions in the first two years, the fewer accidents a person is likely to have in a subsequent period. The net effect of this difficulty is to distort the results of analysis in some unpredictable way. In fact, in the initial statistical analyses, this distortion was so large that negative weights were produced for criminal code offences for which drivers were likely to be suspended. (The implication of a negative weight is that the offence is associated with fewer subsequent accidents). As a result of this finding, it was necessary to remove from the data-set and from subsequent analysis

RELATIONSHIP BETWEEN AGE AND ACCIDENTS

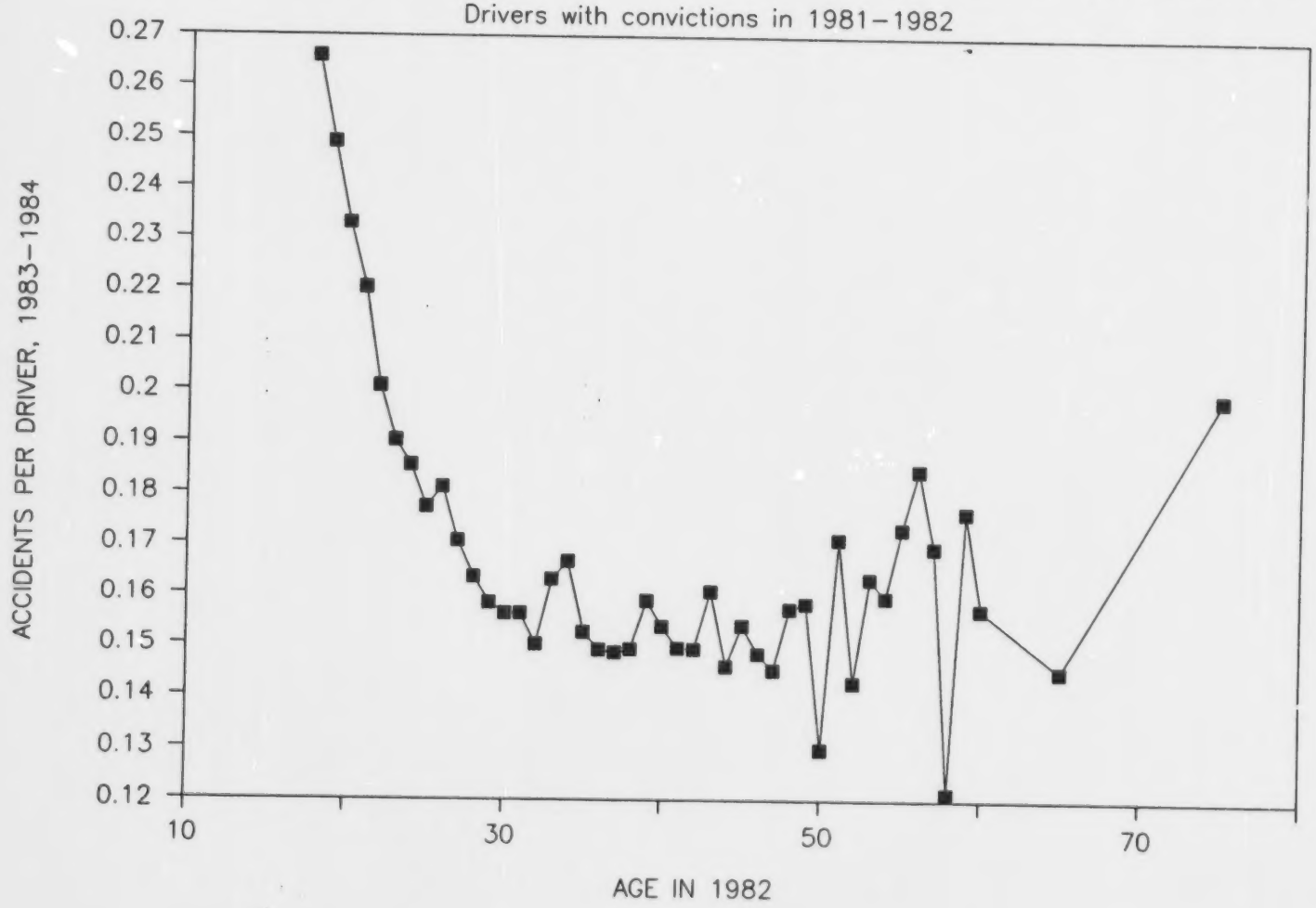


Figure 4

those drivers whose license had been suspended during any time in the period 1981-1984. The result of this decision is that whatever results are obtained in the course of the analysis apply directly only to those drivers whose licenses have not been suspended. The extension to drivers who were suspended under the present demerit point system is therefore an extrapolation.

4.0 ANALYSIS AND RESULTS

The various activities described so far (checking representativeness, conducting consistency checks, the exploratory analysis of data, the selection of conviction categories, determination of age groups and the removal of suspended drivers) are all preliminary to the main activity, namely the establishment of a relationship between information contained in a driver's record and his or her expected number of future accidents. The information used was a driver's gender, age, the count of accidents (at-fault, not-at-fault or total) and the count of convictions in each of the fourteen categories. This information from the first two years was used to estimate "regression weights" which best fitted the accident record in the second two-year period. These regression weights are the relative number of points each conviction category should be assigned for the best prediction of the likelihood of an accident in the second two-year period. Only records of those drivers who had at least one conviction in the first two-year period were used (n ~170,000).

4.1 Schemes and Variants Examined

The MTO was interested in a number of variants, each using different sets of variables. These fall into three categories:

1) models which made use of age and gender information and models which did not;

2) models which assigned points for accidents (with the further distinction between at-fault, not-at-fault and total) and models which did not assign points for accidents;

3) models which assigned different numbers of points for each conviction category and models in which all convictions carried the same weight.

In total, 16 different combinations of variables from the first two-year period were used. For each of these, weights were estimated. Each of these 16 combinations results in a prediction equation which is termed a "model". Table 4 shows which variables were used in each "model". Thus, for example, model B2 calculates the expected number of period-two accidents using the count of period-one convictions in each of the 14 groupings as well as the count of total period-one accidents; model C4 uses driver age and gender, the count of all convictions (without distinction with respect to grouping), the count of at-fault accidents and the count of not-at-fault accidents.

TABLE 4: VARIABLES USED FOR REGRESSION RUNS
(x -- indicates variable used in run)

Run	Age and Sex dummy variables	Variables for each Conviction Group	Variable for total Convictions	Accident variables		
				Total	Fault	At-Fault
A1	x	x				
A2	x	x		x		
A3	x	x			x	
A4	x	x			x	x
B1		x				
B2		x		x		
B3		x			x	
B4		x			x	x
C1	x		x			
C2	x		x	x		
C3	x		x		x	
C4	x		x		x	x
D1			x			
D2			x	x		
D3			x		x	
D4			x		x	x

Half of the models tested (A1-A4, B1-B4) allowed for a different weight to be given to each of 14 conviction types. The other 8 models tested used a weighting of 1 point per conviction, that is, all conviction categories were considered equivalent (models C1-C4, D1-D4).

Half of the models used age and gender as variables, to aid in the prediction of accidents in the second 2 year period (models A1-A4, C1-C4). As will be seen in section 4.4, where the performance of the various models is examined, using age and gender, as well as convictions and accidents greatly improves ability to predict future accidents. This is because there is such a strong correlation between these two variables and accidents. For example we can be sure that young male drivers will have more accidents than other drivers.

One-quarter of the models (A1, B1, C1 and D1) used only information on convictions in the first two year period to predict accidents in the second two year period. One-quarter of the models (A2, B2, C2 and D2) used total accidents in the first two year period, as well as convictions, to predict subsequent accidents. One-quarter (A3, B3, C3 and D3) used only at-fault accidents and convictions for prediction purposes. Finally, one-quarter (A4, B4, C4 and D4) used separate weights for at-fault and not at-fault accidents, as well as convictions (also separately weighted: A4, B4 or one point per conviction: C4, D4) to predict subsequent accidents.

4.2 How a Model Works

Details about the models are given in the full report. Here the essential nature of the model will be illustrated, using Model A2. Consider a female driver, 24 years of age, who in the first period had two speeding convictions, one conviction for failing to yield the right of way and one at-fault accident. Regression weights for Model A2 are shown in Table 5. The base driver for all the models is a 17-20 year old male who is conviction and accident free in the first period, and who is expected to have 0.176 accidents in the second period. From this value one has to subtract 0.061 for being female and 0.039 for being 24 years old and add 0.027 for each speeding conviction and 0.027 for the right-of-way conviction and 0.058 for the accident. On the basis of model A2, we expect this driver to have:

$$0.176 - 0.061 - 0.039 + 2(0.027) + 0.027 + 0.058 = 0.215 \text{ accidents}$$

in the second two-year period.

4.3 The Accident Potential of Ontario Drivers

The models estimate for each driver the number of accidents he or she is expected to have per year in the second two year period. This is not an integer because its nature is that of a "long-term-average". For example an estimate of 0.25 accidents in the second two year period means that driver is expected to have 1 accident in the next 8 years. For brevity, this number is called a driver's accident potential. Of course, not all drivers have the same accident potential; some drive more, some drive less, some take risks, others are more cautious. Prior to examining results for each model, let us examine the diversity of accident potential existing in the population of

TABLE 5: REGRESSION COEFFICIENTS FOR MODEL A2

Description of variable	Estimated Coefficient	Standard Error
Intercept (Male, Age < 21)	0.17630	0.002876
Dummy variable for age 21-25	-0.03856	0.002852
Dummy variable for age 26-30	-0.05650	0.002913
Dummy variable for age 31-35	-0.06006	0.003044
Dummy variable for age 36-40	-0.05202	0.003516
Dummy variable for age 41-50	-0.06013	0.003510
Dummy variable for age 51-60	-0.05774	0.004832
Dummy variable for age > 60	-0.06156	0.007068
Dummy variable for female	-0.06122	0.001735
mm1: Fail to yield, imp. turns, PXO, amber violations, etc	0.02696	0.001753
mm2: Disobey red lights; rail crossing violations.	0.042125	0.002916
mm3: Fail report accid.; Careless driving; Dang. driving;Crim. neg. caus. death	0.023304	0.008613
mm4: Unsafe move;imp. o'taking; Disobey signs	0.06359	0.004959
mm5: Fail to remain; breath test; alcohol; impairment.	0.2444	0.054575
mm6: Seat belt; F.T.C.; parking; divided h'way offences.	0.03221	0.002130
nn1: Minor neglect: no drivers license; permits; insurance, address change	0.02624	0.010463
nn2: Neglect: D/Lic. suspended or not produced; plates, insurance.	0.03366	0.003708
vv1: Minor vehicle neglect: lamps windows obstructed, etc.	0.04954	0.006561
vv2: Vehicle neglect: Unsafe veh., brakes, tires	0.08673	0.009296
vv3: Insecure load.	0.159215	0.027210
vv4: Weights and dimensions offences	0.11074	0.011895
ee1: Environmental offences: noise, fumes.	0.08748	0.011895
mm7: Speeding offences	0.026515	0.000866
Total accidents in period 1	0.05831	0.001580

Ontario drivers. This will reveal how many drivers there are in the population who have a high accident potential. How many of these "high accident potential" drivers will indeed be identified for post-licensing-control under the current and the new weighting models, will be examined later. The number of accidents in the second two-year period was used to compute the mean accident potential (0.055 accidents/year) and its standard deviation (0.055 accidents/year) in the total driver record sample. Details of the method are given in the full report. This information was then used to plot the distribution of accident potential in a population of 5 million Ontario drivers shown in Figure 5.

Using Figure 5, one can establish how many drivers in the population are at each level of accident potential. Thus, for example, it can be seen that about 1.75 million drivers, or 1/3 of the drivers, have an accident potential which is higher than the average. (It is less than 1/2 of the population because there are some drivers with a very high accident potential). Let us now find how many drivers there are with an accident potential which is much higher than the mean for this population. Consider drivers whose accident potential is 3 standard deviations above the mean or higher, that is, an accident potential larger than 0.22 accidents/year. From Figure 5 it can be seen that almost 90,000 drivers have an accident potential of 0.22 or higher.

It is drivers with a high accident potential who are to be identified for post-licensing control. However, it must be noted that a driver may have a high accident potential because of bad driving as well as because of high exposure to driving. If a driver is likely to have x accidents driving y kilometers, the same driver is likely to have 2x accidents driving 2y kilometers. Thus, it is very important to ascertain whether one has identified, through the demerit point system, a driver who drives little but has a high accident potential or one who drives a lot and has a correspondingly high accident potential. The two drivers may require different treatments.

4.4 The performance of the current demerit point system and of the new models

Since the 16 new models were derived using appropriate statistical methods, rather than by weighting each offence according to its perceived seriousness, they should perform better than the current demerit point system. However, all models face the same inherent difficulties as the current demerit point system. Namely, because of the randomness inherent in the process of accident occurrence and the randomness inherent in the process by which drivers acquire convictions, a two-year record is just too short to accurately estimate a driver's accident potential. As will be seen, the new models are an improvement on the current system, but like the current system, still fail to detect many of the high accident potential drivers.

Two measures of performance will be used to judge the quality of a model. The first is the number of period-two accidents which have occurred to 10,000 drivers selected as worst by a model. The model which identifies drivers with the most period-two accidents will be judged as best.

Distribution of Accidental Potential in the Ontario Driver Population

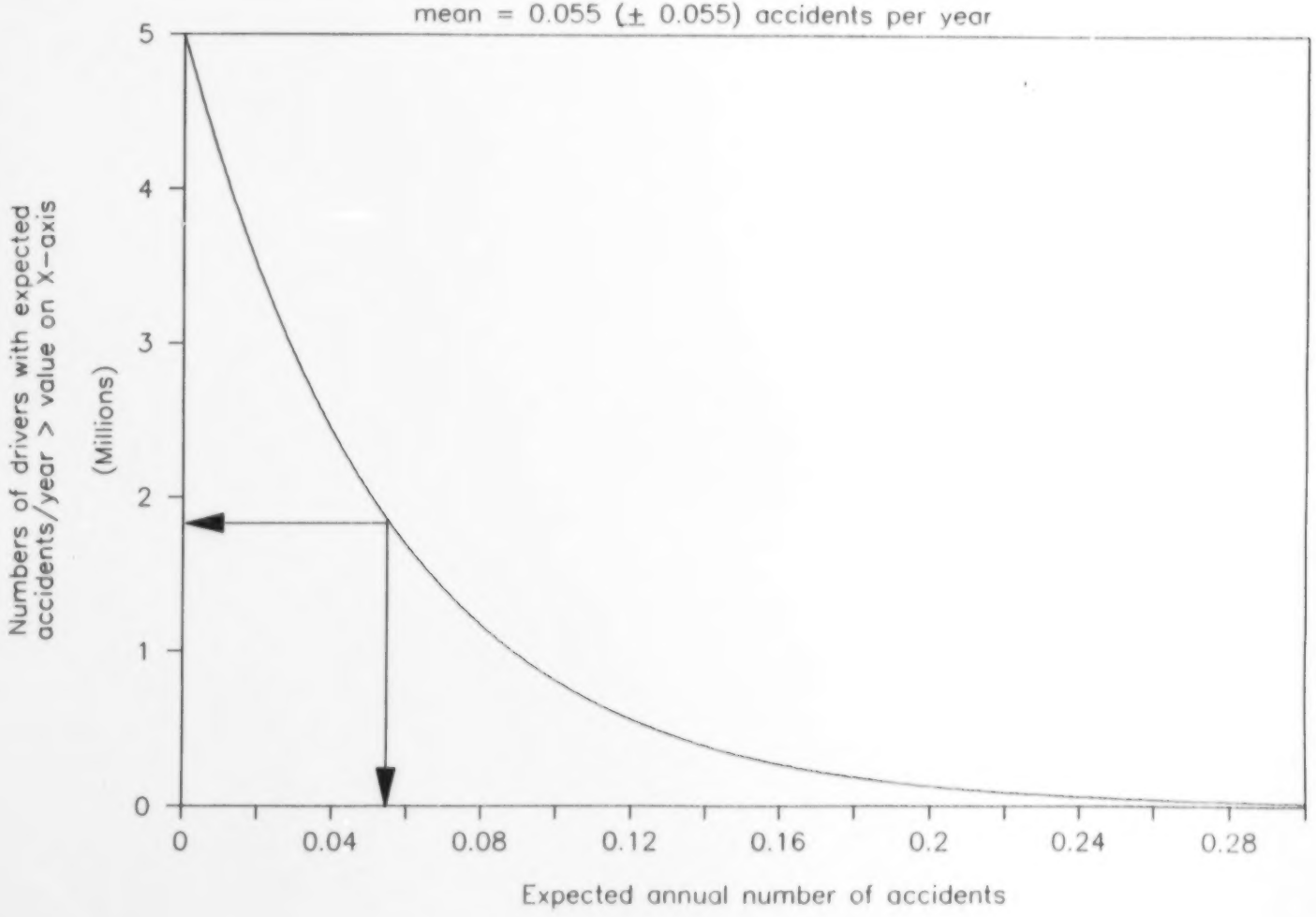


Figure 5

The second measure of performance gives more detailed information. It tells us how many of the drivers estimated to have a high accident potential on the basis of their period-one record are truly high accident potential drivers. These are the "hits". We also learn how many of these drivers have a true accident potential which is lower than the population average. These are the "false alarms". These measures will be explained in detail below.

First Performance Measure: Future Accidents. The first measure of performance is straightforward. Consider, for example, the 10,000 drivers who in the first two-year period had the most demerit points. Checking the accident records of the same drivers, we find that during the second two-year period they had 1452 (180+640+632) accidents per year (see Table 6). Consider now another group of 10,000; this time those who in the first two year period had the most accidents. For this second group, we find they recorded in the subsequent two-year period 1828 (312+756+760) accidents per year. Evidently, it is better to identify drivers by their previous accident record than by previous demerit points. Imagine now that a third group of 10,000 drivers is identified, this time those for whom model A4 estimates the highest accident potential on the basis of their age, gender, as well as convictions and accidents in the first two-year period. This group has 2084 (320+868+896) accidents per year in the second period. Thus, selection by model A4 gives a richer catch than either selection by previous accidents or by the current demerit points.

The performance of the sixteen models, the current demerit point weights and "previous accidents" is described in Table 6. In interpreting these results, one has to keep in mind that the count of accidents is always subject to random fluctuations.

Several conclusions emerge. First, one can do better than to use the current demerit point weights. Second, it is important to make use of the driver's record of accidents. This is already evident from the comparisons of the last two rows. However, it also emerges from the poor performance of models A1,B1,C1 and D1 which do not make use of accident data. In fact, the top 1000 drivers can be well identified by their previous accident record alone. Third, the more drivers are identified, the lesser the "yield". Thus, the top 1000 drivers have an accident rate of ~0.3 accidents/year which is approximately 6 times the population average; for the first 10,000 drivers the average accident rate is ~ 0.2 and so on.

The measure of performance examined so far leaves the impression that the drivers which are identified indeed have an accident potential which is substantially larger than that for the population of all drivers. While this is true for the group "on the average", this group itself may not be a homogeneous one. The second measure of performance by which the quality of the alternative models is to be judged relates to the diversity of accident potential within the group of drivers which these models identify.

Second Performance Measure: Identifying Hits and False Alarms: A weighting scheme is like a net with which we attempt to catch drivers who, based on their 2 year record, are likely to have an unusually high number of accidents in the next 2 years. For the illustration here, consider

Table 6: ACCIDENTS PER YEAR RECORDED BY DRIVERS SELECTED BY VARIOUS MODELS

Model	Drivers estimated by model to be					
	Top 1,000	Next 4,000	Next 5,000	Next 10,000	Next 100,000	Total
A1	188	712	904	1660	13308	16772
A2	324	856	936	1798	14192	18016
A3	276	736	860	1712	13700	17284
A4	320	868	896	1736	14060	17880
B1	212	704	824	1548	12888	16176
B2	.320	832	980	1628	13688	17448
B3	272	748	804	1712	13336	16872
B4	304	876	956	1636	13684	17456
C1	208	744	760	1592	13016	16320
C2	356	808	900	1672	13732	17468
C3	276	780	780	1424	13684	16944
C4	356	804	928	1616	13748	17452
D1	176	748	688	1432	12260	15304
D2	352	824	784	1608	13268	16836
D3	244	756	852	1360	13084	16296
D4	364	840	788	1576	13216	16784
Accs.*	312	756	760	1432	10780	14040
DP**	180	640	632	Not	Available	

* - Drivers with the highest accident counts in period 1 were selected.
** - Drivers with the highest demerit points acquired in period 1.

"unusually high" to be 3 standard deviations above the mean; hat is, we are hoping to identify drivers whose accident potential is larger than 0.22 accidents per year. If we manage to identify such a driver we will call it a "hit". Conversely, if based on the 2 year record we identify, and call in for treatment, a driver whose accident potential is lower than what is the average accident potential in the population, that is 0.05 accidents per year, we will call it a "false alarm".

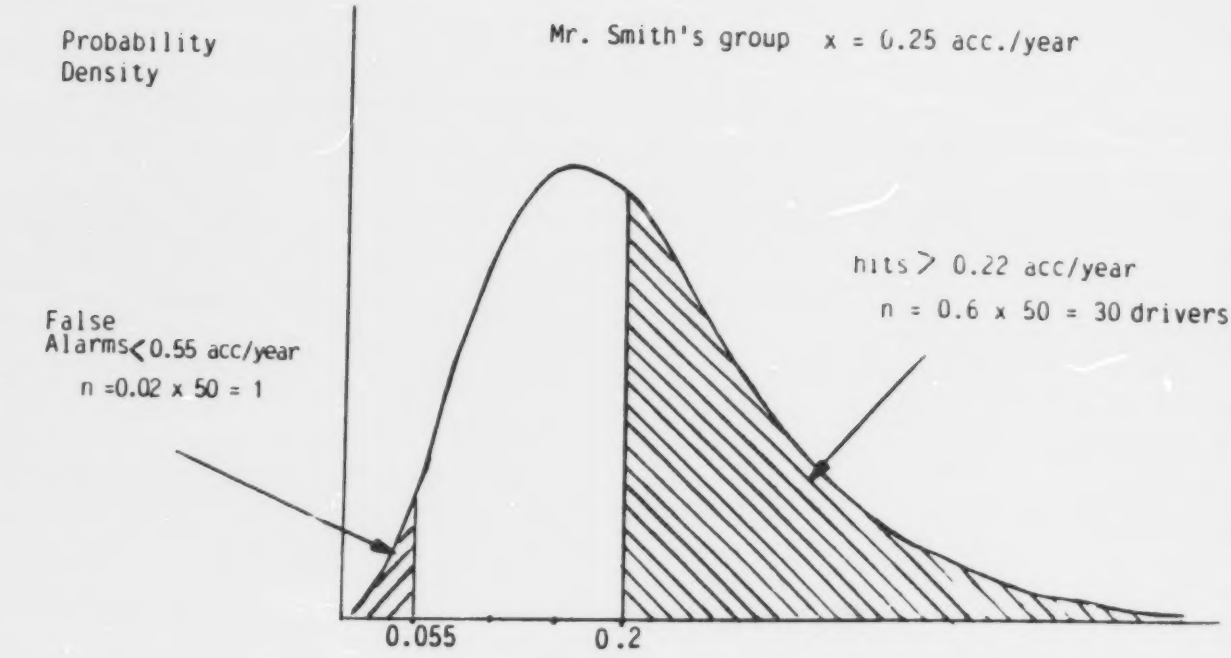
For illustration purposes, models A4 and B4 were used. (The variables which these use to estimate accident potential are shown in Table 4.) Drivers were ranked in terms of accident potential estimated by each model, based on their record during the first two-year period. Table 7 shows the hits (drivers correctly determined to have accident potential larger than 0.22 accidents/year) for the 10,000 drivers from a population of 5 million drivers with the highest accident potential as estimated by each model. Prior to interpreting the results of Table 7, a brief discussion of how the numbers in the table were derived is given. Determining the hits and false alarms for each weighting scheme, or model, is a lengthy and rather complex procedure. It is not simply a matter of counting how many drivers had accidents in the second two year period, and how many did not. An explanation of the process, in general terms, follows.

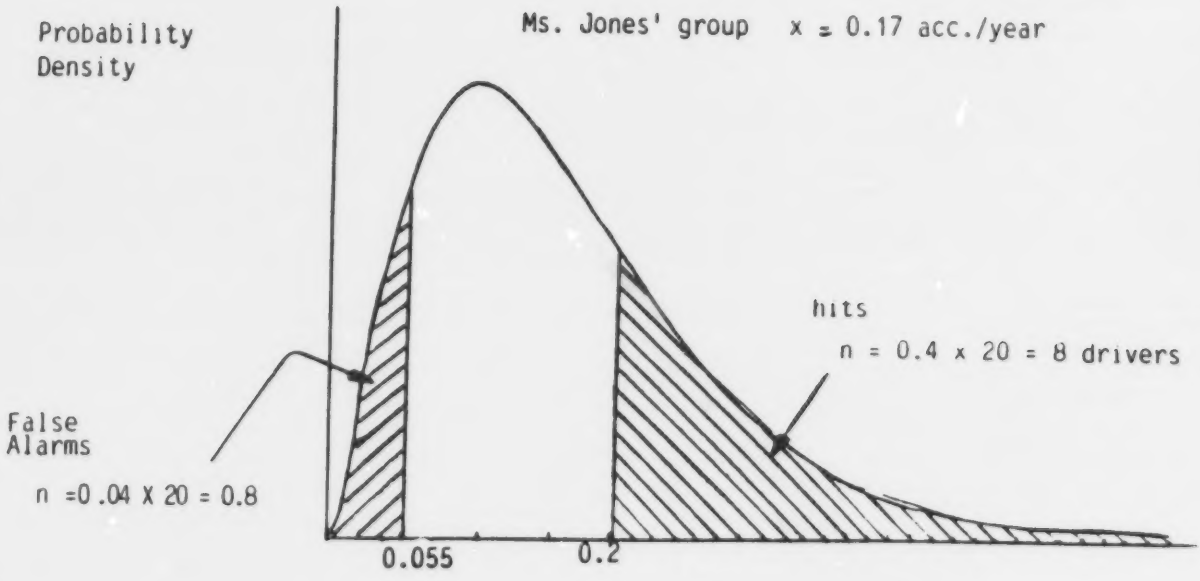
Let us consider one driver out of a sample of 827,995, Mr. Smith. Mr. Smith's record of convictions and accidents in the first two year period is used to estimate his expected number of accidents per year in a subsequent time period, using one of the point schemes or models that have been developed, for example, A4. (See section 4.2 for details about the estimation of expected numbers of accidents using a driver's record and model A2). This process is then repeated for all other drivers. Now let us select those 1000 drivers who are estimated (say, using model A4), to have the highest expected number of accidents. (In the current demerit point system this is equivalent to selecting the 1000 drivers who have the highest number of demerit points). Let us suppose that Mr. Smith is among these. Amongst the 1000 drivers there will be other drivers of the same age group, sex, and record as Mr. Smith. These drivers, though similar to Mr. Smith in these three characteristics, will drive different amounts, in different driving styles and are expected to have different numbers of accidents. For example, some who are in fact good drivers may have been in accidents or received convictions as a result of acts they seldom commit. These are expected to have few accidents in a subsequent period. Others in this group are in fact poor drivers and are expected to have a large number of accidents in the subsequent period. Thus, the population of drivers who are in the same age group, sex and driving record as Mr. Smith in the first two year period, in fact have very different expected numbers of accidents. However, to the model, these people appear to be the same, and thus are assigned the same *estimated expected number of accidents*. The model reflects the diversity of real accident potential in this group, by not only telling us the *average* expected number of accidents, but also the *variance* of expected number of accidents in this group. Thus, for Mr. Smith and drivers of the same age group, sex and record, there is a distribution of expected number of accidents which is estimated by our model. As shown in Figure 6 for this group of drivers, with an average expected number of accidents of 0.25, 60% in reality have an expected number of accidents of 0.22 (hits) or more, while 2% at the

TABLE 7: Figures of merit for Models A4, B4

Drivers estimated by model to be in		Number of drivers expected to have $m > 0.22$		Number of drivers expected to have $m < 0.05$	
		Model A4	Model B4	Model A4	Model B4
the top	1,000	528	541	39	45
the next	4,000	1568	1595	246	289
the next	5,000	1601	1620	390	458
the next	10,000	2679	2657	933	1110
the next	100,000	15987	15291	14198	16928
TOTALS	120,000	22363	21704	15806	18830

Figure 6


Probability Density
Mr. Smith's group x = 0.25 acc./year
hits > 0.22 acc/year
n = 0.6 x 50 = 30 drivers
False Alarms < 0.55 acc/year
n =0.02 x 50 = 1
0.055
0.2



Expected Accidents/Year

Note: These graphs are for the purpose of illustration only and are not meant to accurately represent particular data sets.

other end of the distribution have in reality an expected number of accidents of 0.055 or less (false alarms).

Now consider another driver, Ms. Jones, who belongs to a different group of drivers amongst our population of 1000 drivers with the worst records. In this group, the average expected number of accidents is different from Mr. Smith's group (though it is still high), and the curve of the distribution of actual expected numbers of accidents is different (See Figure 6). Here it may be that 4% of the drivers have a low expected number of accidents of 0.055 or less (we call these *false alarms*), while 40% have an expected number of accidents of 0.22 or more (which we call *hits*).

In order to determine the total number of hits (drivers with an expected number of accidents of 0.22 or more) in the population of 1000 drivers with the worst records, we add the numbers of drivers in each of the various groups (the Mr. Smith group, the Ms. Jones group, etc.) who would have more than 0.22 expected accidents. Let us assume there are 20 drivers in Ms. Jones' group, and 50 drivers in Mr. Smith's group. In the case of the Mr. Smith group, there will be 60% of 20 drivers, or 12 hits. In the case of the Ms. Jones group, there will be 40% of 20 drivers, or 8 hits. By adding up the hits in each of these distributions, we obtain the figure shown in Table 7; that is, 528 hits in the 1000 worst records by model A4, 1568 hits in the next 4000 worst records, etc. Similarly, we can add up the false alarms in each of our distributions.

As Table 7 shows, in a group of the worst 1000 drivers as selected by model A4, there are 528 hits and 39 false alarms. When drivers are selected by model B4, the number of hits increases to 541 and the number of false alarms correspondingly increases to 45. Thus, there is not much to differentiate the two models for the top 1000 drivers, probably because they consist largely of the same persons. In the second 10,000 worst driver records, A4 gives 2679 hits and 933 false alarms while B4 gives 2657 hits and 1110 false alarms. Thus, A4 is not much better than B4 in terms of identifying the hits, but is substantially better in terms of calling in fewer drivers which are better than average.

Note that the yield of hits is above 50% in the first 1000 drivers and only ~15% in drivers ranking between 20,000 and 120,000. Similarly, there are only 3-4% false alarms in the top 1000 drivers and this increases to 14-17% in the last row. Thus, the larger the group identified, the worse the performance of the net.

Now let us consider another aspect of how our net performs. How many of the drivers in the population who have an unusually high accident potential will slip through the holes? Earlier, we showed that there are drivers with bad records who belong to a group with a high average accident potential, but who themselves belong to the end of the distribution where the probability of an accident is low. Similarly, there are many drivers with average or good driving records who belong to groups with low average accident potential, but who themselves are in the upper end of the distribution and have a high accident potential. Such individuals will slip through the net.

Looking again at Figure 5, it can be seen that out of 5 million drivers 90,000 drivers have an accident potential larger than 0.22 accidents/year. Thus, as Table 7 shows, using model A4 to select the 10,000 drivers with the worst records will catch 3698 of the high accident potential drivers; calling in the next 10,000 will identify 2679 more hits. Calling in the next 100,000 will yield another 15,987 hits. Thus, even after those 120,000 drivers of 5 million who have the highest accident potential according to model A4 have been selected for treatment, only 22,363 hits can be expected. Of the 90,000 drivers in the population who have an unusually high accident potential (>0.22) still 67,637 remain still unidentified. One would have to call in the whole driver population before all the hits would be identified.

Table 8 compares performance among the 16 models, and the current demerit point system, in terms of accuracy of prediction for the worst 10,000 drivers identified by each model. It should be noted that, while there always will be considerable overlap between groups of drivers identified by different models, there also will be systematic differences. Thus, for example, the *use of model series A and B will lead to groups which contain more truck drivers than the current system,* simply because the current system does not assign any points for truck weight or truck dimension offences, while models A and B weigh these heavily.

While comparison in terms of hits and false alarms is good for purpose of illustration, it depends on a rather arbitrary definition of what is to be considered an "unusually high" accident potential. A more comprehensive way to characterize the performance of different models is by continuous curves, as shown in Figure 7.

In Figure 7, accident potential is measured on the horizontal axis. The lowest curve represents the current demerit point scheme. In a group of 10,000 drivers who in a population of 5 million have the most demerit points, one can expect to find 2800 who have an accident potential above 0.2 accident per year. The highest curves represent models A4 and B4. In a group of 10,000 drivers who in a population of 5 million have the highest estimated accident potential by model A4, one can expect to find some 4200 drivers who have an accident potential above 0.2 accidents per year. Thus, the higher the curve, the better the net. (Note that accident potential above 0.2 accidents per year is determined here, slightly less than the definition used for a hit of 0.22 accidents per year. Thus for model A4, 3698 were identified as hits vs. 4200 here.)

While the current demerit point scheme is clearly inferior to the new schemes, there is not much to distinguish between the other four models in Figure 6. A4 and B4 give a separate weight to each of 14 convictions and are somewhat better than C4 and D4 which assign the same weight to all convictions. A4 (which uses age and gender information) is slightly better than B4 where accident potential is less than 0.15 accidents/year. B4 is slightly better than A4 for accident potential greater than 0.15 accidents/year.

In Figure 8, we show a second group of 10,000 drivers, those who rank in the 20,000 to 30,000 worst records; hat is, these drivers' records are bad, but not as bad as the 10,000 shown in Figure 7. The difference in the performance of the four models appears to be even less important.

TABLE 8: Figures of merit for 10,000 drivers with highest m's (for each model)

Model	Number of drivers expected to have m > 0.22	m < 0.05
A1	3258	908
A2	3691	676
A3	3449	817
A4	3698	674
B1	3331	1062
B2	3750	806
B3	3516	923
B4	3757	792
C1	2911	1024
C2	3411	756
C3	3147	922
C4	3429	752
D1	2978	1211
D2	3441	909
D3	3155	1101
D4	3451	906
CURRENT DP	1983	1465

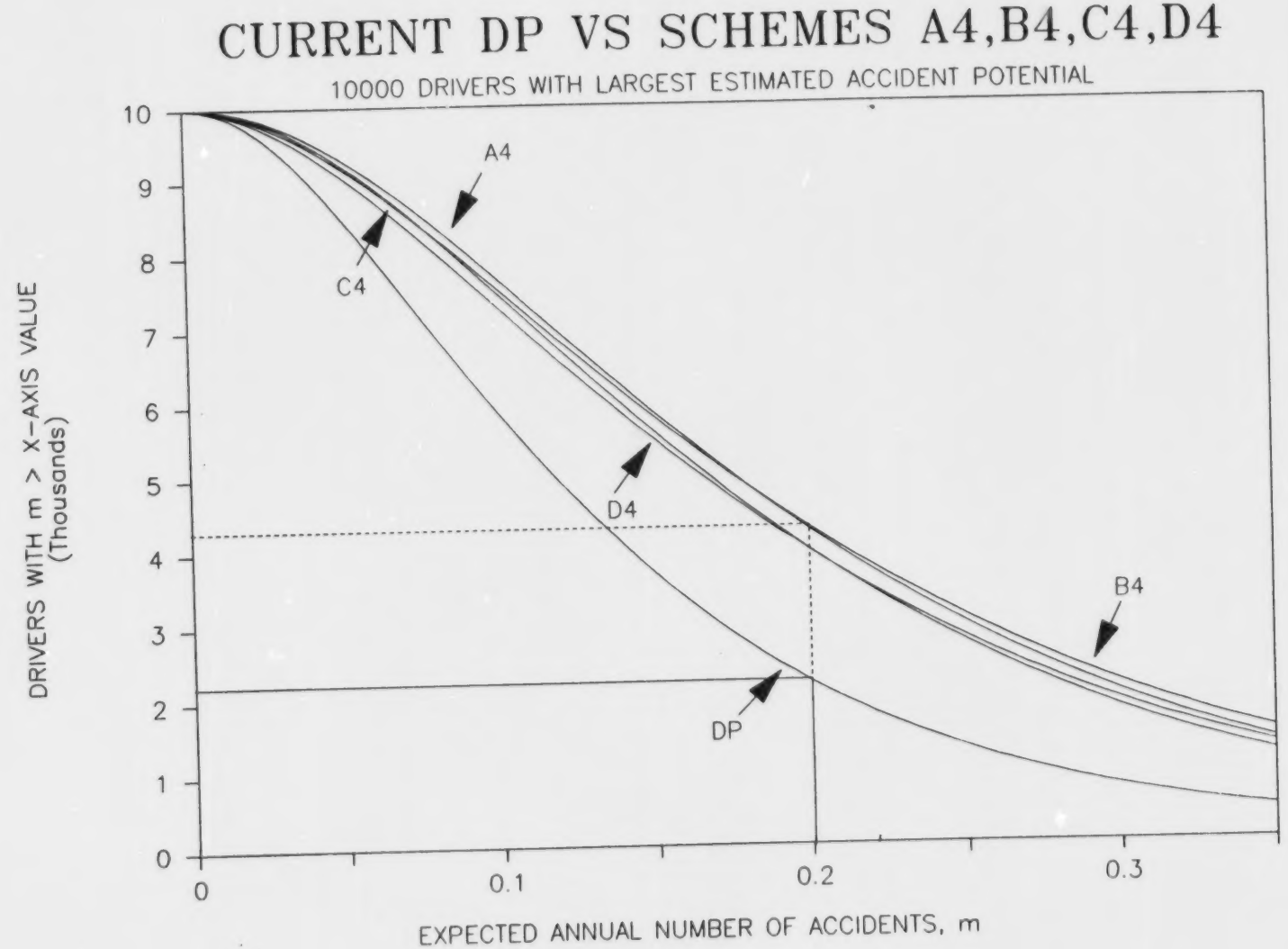

CURRENT DP VS SCHEMES A4,B4,C4,D4
10000 DRIVERS WITH LARGEST ESTIMATED ACCIDENT POTENTIAL
DRIVERS WITH m > X-AXIS VALUE
(Thousands)
EXPECTED ANNUAL NUMBER OF ACCIDENTS, m
A4
B4
C4
D4
DP
0
1
2
3
4
5
6
7
8
9
10
0
0.1
0.2
0.3

Figure 7

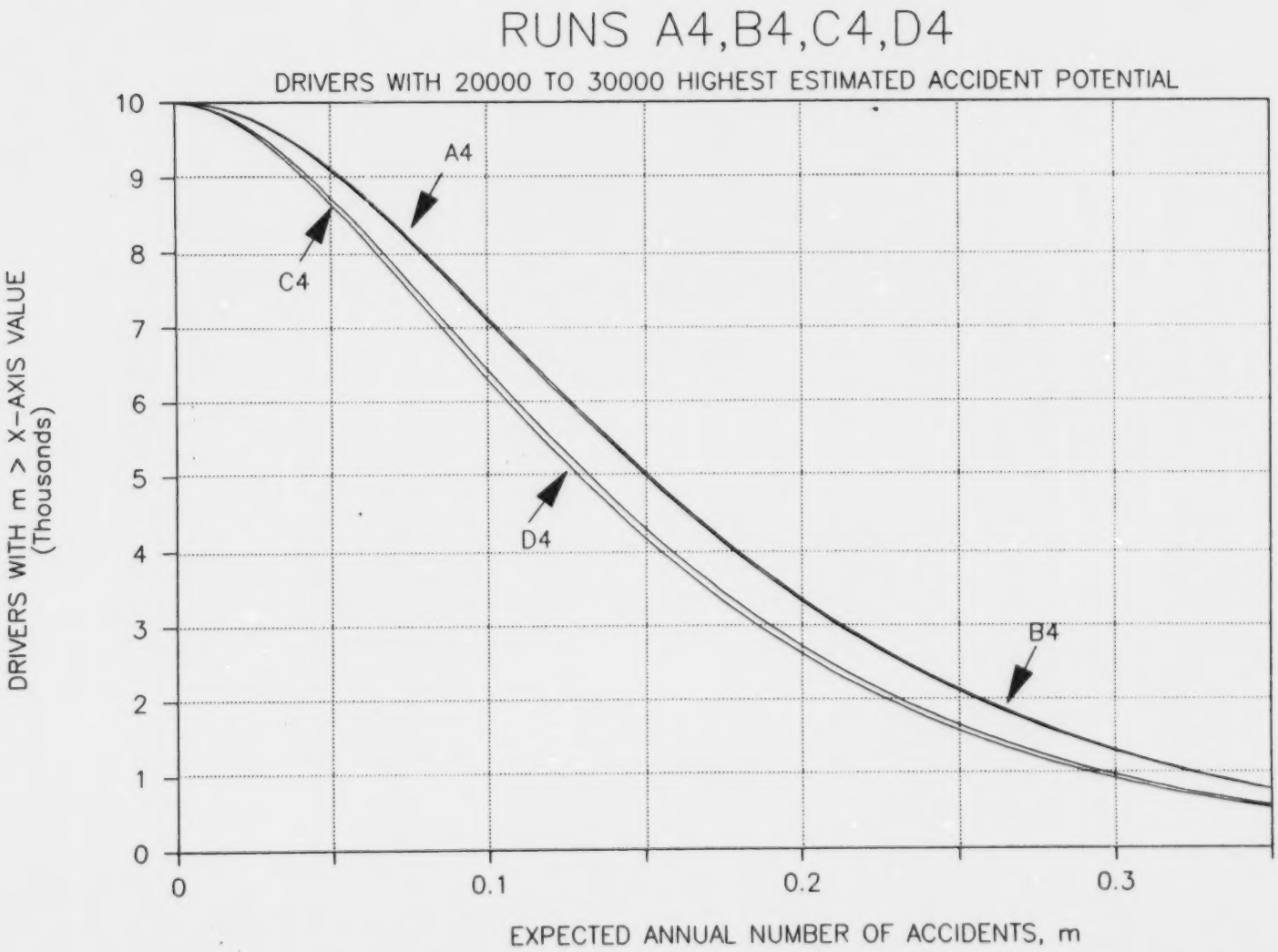

RUNS A4,B4,C4,D4
DRIVERS WITH 20000 TO 30000 HIGHEST ESTIMATED ACCIDENT POTENTIAL
DRIVERS WITH m > X-AXIS VALUE
(Thousands)
10
9
8
7
6
5
4
3
2
1
0
A4
C4
D4
B4
0
0.1
0.2
0.3
EXPECTED ANNUAL NUMBER OF ACCIDENTS, m

Figure 8

5.0 SUMMARY AND CONCLUSIONS

The purpose of this work was to use a driver's record of convictions and accidents in order to predict, as well as possible, which drivers, based on their past record, are most likely to have an accident in the near future.

Driver Record Sample. A sample of 827,995 records of drivers licensed to drive in Ontario during 1981-1985 has been examined. Each driver record contained information about the driver's gender, age, and details of accidents, convictions, demerit points and suspensions.

Preparatory Analysis. In preparation for analysis, the many hundreds of offence types had to be grouped into a manageable number of categories. We first put in the same group offences which were similar in nature, and then consolidated those offences which were associated with a similar average number of accidents. It turned out that, for example, drivers who in one year had a single speeding conviction had fewer accidents in the remaining three years than other drivers who had a single conviction in that year for a relatively minor offence, such as a missing lamp. This finding may be initially puzzling but, on reflection, helps to interpret later results correctly. It arises partly because not all illegal behaviours lead to convictions at the same rate, partly because some offenses are specific to truck drivers who drive 10-20 times as much as car drivers and tend to have proportionately more accidents, and partly because behaviour which results in a fairly innocuous offense, such as a noisy muffler, may be of the type which also leads to accidents. Therefore, neither the ratio of accidents to convictions nor the "weights" which are later attached to particular offences are an indication of the gravity of that offence.

Drivers whose license is suspended will curtail their driving to some extent. This is why, during the period of suspension, one should expect some reduction in the number of accidents in which they are involved. However, the extent to which driving is curtailed is unknown. To assume that suspended drivers stop driving would introduce a bias into the analysis; assuming that they continue to drive would cause another bias. To protect the integrity of the results, drivers who were suspended had to be removed from the analysis. Therefore, our conclusions apply directly only to drivers who have not been suspended. The extension to drivers who were suspended under the present demerit point system is therefore an extrapolation.

Alternate Weighting Schemes. We have examined 16 models to estimate the expected number of future accidents for a driver based on age, gender, convictions and accidents. The models differ from each other in the information which is used. Some make use of age and gender, others do not; in some, each of 14 types of convictions is given a different weight, in others all convictions have the same weight; in some, at-fault accidents are counted separately from not-at-fault accidents, in some they are lumped together. All 16 models have a common structure, that of a "weighted sum" of convictions, and in some models, accidents. The basis is always a 17-20 year old male with no accidents or convictions in the first two years. Thus, in one of the models, these drivers are expected to have 0.176 accidents in the second two year period. For being female one subtracts 0.061, for being 24 years old one subtracts 0.039, for having two speeding convictions one adds 2 x 0.027, for having an accident one adds 0.058, and so on. The result is an

estimate of the accident potential of a driver with specific age, gender and record of convictions and accidents. We also estimate the distribution of accident potential in the population of drivers with given traits. We make use of this distribution to describe the performance of alternative models.

Performance of Alternate Schemes and the Current Demerit System. Not all drivers have the same expected number of accidents per year. Some drive more, some less, many drivers are prudent, some take unwise risks. Based on the accident data, we show how many drivers in Ontario have what expected number of accidents. Thus, for example, of 5 million drivers, some 90,000 have an expected number of accidents three standard deviations above the average for the population. It is these "high-accident-potential" drivers which a demerit point scheme aims to identify.

Two measures of performance were used to judge the quality of a model. The first measure of performance is the number of second period accidents in a group of drivers (of fixed size) identified on the basis of their first-period record. Thus, we find that those 10,000 drivers who in the first two-year period had the most demerit points recorded during the second two-year period 1452 accidents per year. Those 10,000 drivers who in the first two year period had the most accidents recorded in the subsequent two-year period 1828 accidents per year. A third group of 10,000 drivers, those who by model A2 have the highest accident potential when calculated on the basis of first period data, recorded 2116 accidents per year in the second period. Thus, selection by model A2 gives a richer catch than either selection by previous accidents or by the current demerit points. On this score, model A2 seems to perform best.

Several conclusions emerge. First, one can do better than to use the current demerit point weights. Second, it is important to make use of the driver's record of accidents. Third, the more drivers are identified, the lesser the "yield". Thus, the top 1000 drivers have an accident rate of ~0.3 accidents/year which is approximately 6 times the population average; for the first 10,000 drivers the average accident rate is ~ 0.2 and so on.

The second measure of performance tells us about the variation in the group of drivers, identified on the basis of their first period record as having high accident potential. Because of the inadequacy of the data available (i.e. short driving record coupled with the random nature of accidents and convictions) some of the identified drivers will be true hits and have high accident potential; others will be false alarms who in fact have lower than average accident potential.

A two year record of convictions and accidents is just too short to estimate a driver's expected number of accidents with accuracy. This is because a given two year period may not truly reflect a driver's typical driving behaviour. This is why some of those identified by the model as having the highest expected number of accidents, turn out in reality to be just average drivers. Conversely, this is why most high accident potential drivers may not have, in two years, a record which identifies them as such. We showed, for example, that of the 10,000 drivers who by the "richest" model (A4) were estimated to have the highest expected number of accidents, 3698 have an accident potential in excess of three standard deviations above the mean for the population. At the same time, 674 of those 10,000 were average

drivers or better. Were we to identify a different group of 10,000 drivers, those who in a two year period had the most demerit points (using the current scheme), we would have found 1783 drivers with "high-accident-potential" (instead of 3698 by model A4) and 1465 drivers who had a lower than average accident potential (instead of 674 by model A4). Thus, the model A4 is considerably better than the current demerit point scheme. It identifies more high-accident-potential drivers and fewer drivers who are better-than-average.

Since A4 uses age and gender data, it is not fair to compare it to the current demerit point scheme. A fair comparison is with model B1 (no age, gender or accident information used) or B2(accident information used but no age or gender). The performance of these three models for the 10,000 drivers with highest estimated accident potential is shown below:

	Hits >0.22	False alarms <0.055 acc./yr.
Current Demerit Points	1783	1465
B1	3331	1062
B2	3750	806

Comparison of Equal Point Schemes and Separate Weight Schemes. Finally, it may be of interest to note that little is gained by giving different numbers of points to different offences. Model D1 uses simply one-point-per-conviction and no accident data; model D2 uses one-point-per-conviction and 1.88 points per accident. Model D2 is only slightly worse than model B2 which assigns different numbers of points to each of fourteen offence classes.

	Hits >0.22	False alarms <0.055 acc./yr.
Current Demerit Points	1783	1465
D1	2978	1211
D2	3441	909

In summary, if the purpose of a demerit point system is to identify drivers who are most likely to have an accident, the scheme used now is not as efficient as alternative schemes would be. Even by giving equal weights to all convictions and a weight of ~2 to an accident (D2), one can do much better. It is important to use data about accident involvement, but it does not pay to differentiate between at-fault and not-at-fault accidents.

For the top 5000 (or so) drivers , it makes little difference whether separate weights are assigned to different offences. However, for the next 100,000 drivers, separate weights improve performance (in terms of predicting the number of future accidents) by a few percentage points. Separate weights help to increase the number of hits and to reduce the number of false alarms by some 10%.

Comparison of Schemes Using Age and Gender and Schemes That Do Not. For the top 5000 or so drivers, the inclusion of age and gender information appears to be unimportant. For the next 100,000 drivers consideration of age and gender improves performance (in terms of predicting future accidents) by a few percentage points. Consideration of age and gender does not seem to increase the number of hits but it helps in reducing the number of false alarms by some 10%.

Limitations of any Demerit Point Scheme. With all this, one has to keep in mind that if only a few drivers are identified (say about 10,000) 30-40% of those will be genuine high-accident-potential drivers and only 6-10% will be falsely identified better-than-average drivers. However, only 3% of all high-accident-potential drivers in the population will be in this group of 10,000 drivers. It does not help much to increase the size of the group because performance deteriorates with size. Thus, in a group of 120,000 drivers, only ~19% are genuinely high-accident-potential, whereas 13-16% are falsely identified. Even when as many as 120,000 drivers are identified by the richest model, only 22,363 of the 90,000 "high-accident-potential" drivers are caught in the net.

Further Improvements. The performance of models for the estimation of a person's accident potential can be further improved. We think that consideration should be given to a system which nearly continuously tracks a person's accident potential. If during a certain period of time (measured in weeks) the driver did not acquire a conviction and was not involved in an accident, his or her estimated accident potential would be revised downward. If during that period of time, convictions or accidents were recorded, the estimated accident potential would be correspondingly revised upward. A person's aging, the general accident trend and seasonal variation would also be reflected in these revisions. In this manner, a person's current estimated accident potential could be made a reflection of his or her entire driving history. In such a scheme, there is no need to specify an arbitrary period of time after which points are forgiven.

In the models developed so far, involvement in an accident adds a fixed amount to a driver's accident potential. Under the newly suggested scheme, an accident for a person with an already high accident potential would be weighed more heavily. In general, a "revision" scheme of this nature relies on solid mathematical logic and is expected to perform better than other possible schemes.